

ANALYTICAL AND NUMERICAL METHODS
FOR DETECTION OF
A CHANGE IN DISTRIBUTIONS

SAOWANIT SUKPARUNGSEE

A thesis submitted for the degree of
Doctor of Philosophy
in Mathematical Sciences

Department of Mathematical Sciences, Faculty of Science
University of Technology, Sydney
December 2008

Principal Supervisor: Professor Alexander Novikov



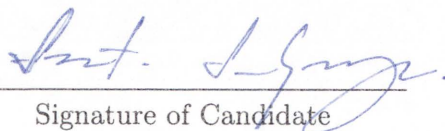
UNIVERSITY OF TECHNOLOGY, SYDNEY
CERTIFICATE OF AUTHORSHIP/ORIGINALITY

Date: Dec 2008

Author: Saowanit Sukparungsee
Title: Analytical and Numerical Methods for Detection of a Change
in Distributions
Department: Mathematical Sciences
Degree: Ph.D. Convocation: Mar Year: 2005

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.



Signature of Candidate

Dedication

*To Dad & Mom
and Family*

Acknowledgments

I would like first to express my deepest gratitude to my advisor, Professor Dr. Alexander Novikov, for introducing me to the wonderful world of martingales, and for his constant support and guidance during my Ph.D. study and thesis preparation. It has been a great pleasure to work under his supervision. I would also like to thank, Dr. Nino Kordazakhia from Macquarie University for all her support.

I am very grateful to Professor Dr. Tony Moon, not only for his invaluable guidance and support throughout my Ph.D. study at UTS but also for their great care and help during my staying in Australia.

I also wish to express my grateful appreciation to all faculty members in the Department of Mathematical Science at UTS, especially Dr. Tim Langtry, Dr. Yakov Zinder and Dr. Nicolae Nicorovici. I can never forget their kindness in assisting me with my Ph.D. study. The warmth of all the staffs and post-graduate students, in particular Gabrile Mititelu here at UTS will live in my heart.

Special thanks go to Dr. Elvin James Moors, Lectures at the Department of Mathematics, King Mongkut's University of Technology, North Bangkok (KMUTNB), Thailand for his careful proof-reading of my thesis and for the helpful suggestions he provided, resulting in substantial improvements to the language of my thesis.

I would like to express my gratitude to the Government of Thailand and the University of Technology, Sydney for their financial support during my study at UTS.

I am also grateful for the generous support from Department of Applied Statistics, KMUTNB, Thailand, where I have worked and developed my academic mental. Thank for giving me the opportunity to study in Australia.

Last, but not least, I would like to thank my mom and family, for their unconditional love, support and patience. To my special friend, Mus, thank you for being there. Finally, I want to dedicate this thesis to my late father, Chanin Sukparungsee, who always called me Doctor Kik in his lifetime.

Table of Contents

Table of Contents	v
List of Figures	vii
List of Tables	viii
List of Notation and Abbreviations	ix
List of Publications and Conference Presentations	x
Abstract	xi
1 Introduction	1
1.1 Formal Setting of Problem	2
1.2 Research Objectives and Contributions	4
1.3 Scope of Study	5
1.4 Thesis Structure	6
2 Statistical Process Control Charts for the Detection of Changes in Pa-	
rameters	7
2.1 The Properties of Control Charts	7
2.2 Shewhart Chart	9
2.3 CUSUM Chart	10
2.4 EWMA Chart	11
2.5 Discussion	13
3 Methods for Finding Analytical and Numerical Approximations for ARL	
and AD	14
3.1 The Markov Chain Approach	15
3.2 Integral Equations	20
3.3 Monte Carlo Simulations	21
3.4 The Martingale Approach	24
3.5 Discussion	37

4 EWMA Chart for Detection of a Change in a Gaussian Distribution 38

4.1 The Derivation of Closed-form Formulas for Gaussian EWMA 38

4.2 Corrected Approximation 44

4.3 Numerical Results 50

4.4 Choices of Optimal Parameters of Gaussian EWMA Designs 55

5 An Adaptation and Expansion of the Martingale Technique to non-Gaussian Distributions 58

5.1 EWMA Chart for Poisson Distribution 59

5.2 Numerical Results 62

5.3 Choices of Optimal Parameters of Poisson EWMA Designs 65

5.4 EWMA Chart for Bernoulli Distribution 68

5.5 Numerical Results 71

5.6 Choices of Optimal Parameters of Bernoulli EWMA Designs 72

6 Conclusion and Recommendations for Further Research 75

6.1 Overall Conclusion 75

6.2 Recommendations for Further Research 76

A Mathematica® codes for Simulation of Sequentially Stochastic Processes 78

A.1 Gaussian Distribution 79

A.2 Poisson Distribution 81

A.3 Bernoulli Distribution 83

B Mathematica® Codes for Calculation of Closed-form Formulas 85

B.1 Procedure for obtaining the optimal parameter 85

Bibliography 96

List of Figures

1.1	Sample of control chart	2
2.1	A typical Shewhart chart	9
3.1	In-control region divided into n subintervals to form n in-control states of the Markov Chain	19
4.1	Comparison of lower-bounds, first approximations and simulations of ARL: Gaussian case	48
4.2	AD_2 for different magnitudes of change: two-sided Gaussian case	51
4.3	Comparison of AD_1 simulated by EWMA and by CUSUM for a one-sided Gaussian EWMA chart	55
5.1	Comparison of lower-bounds, approximations and simulations of ARL_1 : Poisson case	62
5.2	Comparison of AD_1 simulated by EWMA and by CUSUM for a one-sided Poisson EWMA chart	65
5.3	AD_1 for different magnitudes of change: Poisson case	67
5.4	Comparison of lower-bounds, upper-bounds, approximations and simulations of ARL_1 : Bernoulli case	70
5.5	Comparison of AD_1 simulated by EWMA and by CUSUM for a one-sided Bernoulli EWMA chart	72
5.6	AD_1 for different magnitudes of change: Bernoulli case	74

List of Tables

- 4.1 Comparison of numerical results of ARL_1 between lower-bounds, first approximation and MC for one-sided Gaussian EWMA 47
- 4.2 Comparison of numerical results of ARL_2 between lower-bounds, first approximation and MC for two-sided Gaussian EWMA 47
- 4.3 Comparison of numerical results of AD_1 between lower-bounds, first approximation and MC for one-sided Gaussian EWMA 48
- 4.4 Comparison of numerical results of AD_2 between lower-bounds, first approximation and MC for two-sided Gaussian EWMA 48
- 4.5 Comparison of ARL_2 results from our closed-form formulas with other methods for two-sided Gaussian EWMA 52
- 4.6 Comparison of optimal parameter values and AD_2^* for two-sided Gaussian EWMA 53
- 4.7 Comparison Monte Carlo simulations of ARL_1 and AD_1 for one-sided CUSUM and EWMA charts: Gaussian case 54
- 4.8 Optimal parameter values and AD_1^* of a one-sided Gaussian EWMA 56
- 4.9 Optimal parameter values and AD_2^* of a two-sided Gaussian EWMA 57

- 5.1 Comparison of numerical results of ARL_1 between lower-bounds, first approximation, second approximation and MC for Poisson EWMA 62
- 5.2 Comparison of the numerical ARL_1 from the martingale approach with the numerical ARL_1 from the MCA and Monte Carlo simulations. 63
- 5.3 Comparison Monte Carlo simulations of ARL_1 and AD_1 for one-sided CUSUM and EWMA charts: Poisson case 64
- 5.4 Optimal parameter values and AD_1^* of a one-sided Poisson EWMA 66
- 5.5 Comparison of numerical results of ARL_1 between lower-bounds, upper-bounds, first approximation, second approximation and MC for Bernoulli EWMA 70
- 5.6 Comparison of the approximations with Monte Carlo simulations for the Bernoulli EWMA 71
- 5.7 Comparison Monte Carlo simulations of ARL_1 and AD_1 for one-sided CUSUM and EWMA charts: Bernoulli case 72
- 5.8 Optimal parameter values and AD_1^* for a one-sided Bernoulli EWMA . . . 73

List of Notation and Abbreviations

Notation and abbreviations are used in this thesis are:

ARL	Average Run Length
AD	Average Delay
EWMA	Exponentially Weighted Moving Average
CUSUM	Cumulative Sum
MCA	Markov Chain Approach
IE	Integral Equations
MC	Monte Carlo Simulations
H	control limit/boundary
θ	moment of change-point
τ	stopping time/alarm time
α_0	in-control parameter
α	out-of-control parameter
λ	weighted smoothing parameter in EWMA statistic, $\lambda \in (0, 1)$
χ	overshoot
φ	characteristic function
ψ	cumulant generating function (logarithm of the moment generating function)

List of Publications and Conference Presentations

Parts of the work presented in this thesis have been previously published as:

1. Sukparungsee, Saowanit, & Novikov, Alexander. 2006. On EWMA Procedure for Detection of a Change in Observations via Martingale Approach. In *Proceedings of the KMITL International Conference on Science and Applied Science, 8-10 March, Bangkok, Thailand (KSAS 2006)*.
2. Sukparungsee, Saowanit, & Novikov, Alexander. 2006. On EWMA Procedure for Detection of a Change in Observations via Martingale Approach. *KMITL Science Journal; An International Journal of Science and Applied Science*, 6(2a): 373-380.
3. Sukparungsee, Saowanit, & Novikov, Alexander. 2007. Analytical Approximations for Average Run Lengths in EWMA Charts in Case of Light-tailed Distributions. In *Proceedings of International Conference of Mathematical Sciences, 28-29 November, Bangi-Putrajaya, Malaysia (ICMS 2007)*.
4. Sukparungsee, Saowanit, & Novikov, Alexander. 2008. Analytical Approximations for Detection of a Change-Point in Case of Light-tailed Distributions. *Journal of Quality Measurement and Analysis*, 4(2): 49-56.

Abstract

This thesis aims to derive analytical approximations and numerical algorithms for analysis and design of control charts used for detection of changes in distributions. In particular, we present a new analytical approach for evaluating of characteristics of the “Exponentially Weighted Moving Average (EWMA)” procedure in the case of Gaussian and non-Gaussian distributions with light-tails.

The main characteristics of a control chart are the mean of false alarm time or, Average Run Length (ARL), and the mean of delay for true alarm time or, Average Delay time (AD). ARL should be sufficiently large when the process is in-control and AD should be small when the process is out-of-control. Traditional methods for numerical evaluation ARL and AD are the Monte Carlo simulations (MC), Markov Chain Approach (MCA) and Integral Equations method (IE). These methods have the following essential drawbacks: the crude MC is very time consuming and difficult to use for finding optimal designs; MCA requires matrix inversions and, in general, is slowly convergent; IE requires intensive programming even for the case of Gaussian observations.

In this thesis we develop an approach based on a combination of the martingale technique and Monte Carlo simulations. With the use of a popular symbolic/numerical software Mathematica[®], this new approach allows to obtain accurate procedures for finding the optimal weights, alarm boundaries and approximations ARL and AD for the case of Gaussian observations. Further, we show that our approach can be used also for non-Gaussian distributions with light-tails and, in particular, for Poisson and Bernoulli distributions.